MITTEILUNG 108

OPTICAL FLOW

Bernd Neumann

IfI-HH-M-108

Juni 1983

INHALT

Diese Mitteilung ist eine Ausarbeitung von einführenden Bemerkungen zum Themenbereich "Optischer Fluß" auf der Tagung "Interdisciplinary Workshop on Motion: Representation and Perception", die am 3.-5. April in Toronto, Kanada, stattgefunden hat. Im ersten Teil wird diskutiert, welche Konsequenzen sich für die Berechnung und das Auswerten von optischem Fluß aus unterschiedlichen Annahmen über die sichtbare Welt ergeben. Insbesondere werden Spezialfälle unterschieden, bei denen sich Objekte oder Beobachter in eingeschränkter Bewegung befinden. Der zweite Teil stellt die wichtigsten bisher verfolgten Vorgehensweisen bei der Berechnung von Flußfeldern gegenüber. Dabei wird zwischen einer lokalen Verschiebungsberechnung und einer darauffolgenden Integration zu einem global konsistenten Feld unterschieden. Im dritten Teil werden einige Möglichkeiten zur Auswertung von optischem Fluß genannt und die dabei problematischen Aspekte erörtert.

OPTICAL FLOW

Bernd Neumann
University of Hamburg
Schlüterstraße 70
D-2000 Hamburg 13

1. INTRODUCTION

A large number of contributions to this workshop is concerned with computing or making use of optical flow. This is the term now commonly used for an intermediate representation of time-varying imagery where each pixel is assigned a velocity vector describing its temporal displacement in the image plane or - for human vision - in the retinal field. Optical flow can be consciously experienced by human observers (e.g. when travelling in a car) and has early been recognized as a valuable source of information pertaining to the motion and 3D characteristics of a scene (GIBSON 50). Thorough quantitative analyses, however, have only become available during the last five years, when an increasing number of vision researchers turned to motion problems. As can be seen from this workshop, interesting results on how to exploit optical flow are still being uncovered.

Before making use of optical flow it must be computed - unfortunately. As it turns out, no computational theory has get been offered which promises satisfactory results for unrestricted real-world images. Nevertheless, considerable progress has been made in certain restricted situations. This is also documented by several contributions to this workshop. In this introductory survey I shall try to point out the major differences in the approaches taken so far.

Fig. 1 gives a rough sketch of the representations and the processing connected with optical flow. Much of the variety of the research contributions is due to certain assumptions about the visual world. These will be discussed in the following section. The visual world is projected yielding intensity arrays from which optical flow computation per se proceeds. Three rather distinct directions of processing have been proposed. As a first possibility, optical flow is directly computed from the intensity array. The result is usually a dense flow field. Alternately, descriptive elements like prominent points or edges may be computed first. Points usually give rise to a sparse flow field after correspondence is established. Edges lead to a quite different flow compentation due to the remaining degree of freedom. In section 3 these distinctions are elaborated in some more detail. Finally, I shall briefly review ways of extracting useful information from optical flow.

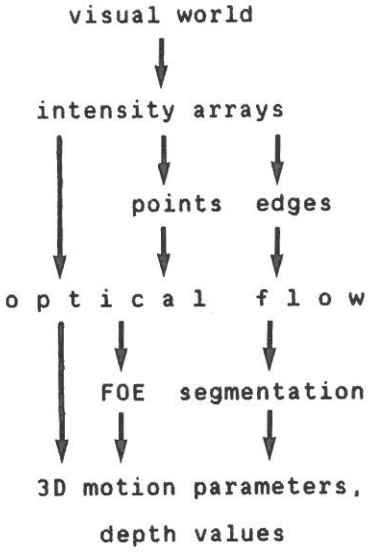

Fig. 1: Alternate ways of computing and exploiting optical flow.

2. ASSUMPTIONS ABOUT THE VISUAL WORLD

Intensity arrays are the result of a physical process which is known to involve an <u>observer</u> (i.e. an optical sensor), <u>objects</u> of the real world, and sources of <u>illumination.</u> In this section some useful restrictions or assumptions regarding these three components will be pointed out. Depending on the choice of restrictions different optical flow problems requiring different solutions may arise.

OBSERVER

For a given scene the complexity of the optical flow field may depend heavily on the motion characteristics of the observer. The following cases have been distinguished so far:

a) stationary
b) translating
c) rotating
d) translating + rotating

While a) is only interesting in connection with moving objects, the other cases are also interesting in a stationary environment. Note that a moving observer in a stationary world is usually not equivalent to a stationary observer facing a single moving object, since the latter case usually also involves stationary background. The distinction between pure translation, pure rotation and the general case is very important for extracting 3D information from optical flow. As shown by PRADZDNY 80 as well as NAGEL and NEUMANN 81, rotation (about the optical center) does not provide depth information. It only complicates the nice properties of optical flow generated by a translating observer. In fact, several approaches to 3D reconstruction from optical flow attempt to "derotate" the images before entering the 3D analysis proper (RIEGER and LAWTON 83). Unfortunately, the full complexity involving both rotation and translation appears to be required for the important case of a translating human observer, since his/her gaze is generally held fixed on a point not straight ahead.

Several subcases of specialized observer motion can also be distinguished. For example, if the direction of an observer translating via stationary environment is known, the focus of expansion (FOE) and hence the direction of the optical flow vectors can be easily determined (JAIN 83). Certain simplifications also arise if the rotation axis is assumed fixed or perpendicular to the direction of translation (BONDE 79).

OBJECTS

There are numerous assumptions and restrictions which may be imposed on the objects of a scene. The following is a list compiled from research dealing with special cases.

a) single / multiple
b) 3D / flat
c) rigid / deforming
d) independent / jointed
e) translating
f) rotating
g) translating + rotating

While most of the difficulties of optical flow computation and subsequent 3D analysis can be studied with a single moving object (particularly if it moves in a non-uniform stationary environment), it is also useful to consider multiple object scenes because of the more complicated segmentation. Single motion concepts like the FOE may sometimes be carried over as shown by JERIAN and JAIN 83.

Another special case arises if objects are essentially flat, e.g. the earth viewed from high altitude. Optical flow computation is far less involved in this case because depth discontinuities do not occur. The mathematical analysis is also easier (TSAI and HUANG 81+82). It is interesting that Gibson considered the approach to a (not necessarily flat) surface fundamental for human spatial perception (GIBSON 50).

Objects need not be rigid. The assumption of rigidity is an essential simplification for a quantitative motion analysis - it amounts to assigning identical 3D motion parameters to all parts of an object. The assumption has not get been brought to bear an optical flow computation, however, which is usually conceived as a process independent of subsequent 3D interpretations. In fact, the 2D smoothness criterion which has been proposed by HORN and SCHUNCK 81 gives rise to nonrigid flow vectors if applied to certain rigidly moving edge shapes (HILDRETH 83). Instead of complete rigidity one may also postulate partial rigidity, in particular jointed motion. Several results pertaining to the analysis of this type of motion are available (WEBB and AGGARWAL 80), all of which assume that the 2D joint positions can somehow be recovered from the flow field. It is not yet clear how this should be done.

Assumptions about the motion characteristics of objects play a similar role for motion analysis as assumptions about observer motion. Depending on the type of restriction, simplified motion equations apply and less complex procedures can be employed. When considering pure rotation it is important to be explicit about the position of the axis. While observer rotation is naturally described about an axis through the optical center, object rotation is sometimes specified w.r.t. other reference systems, e.g. an axis through the object's centroid. It is well known that for a given motion the translation vector (and not the rotation matrix) depends on this choice of reference, hence motion may have a translation component in the object-centered reference system but may be pure rotation if the axis is taken through the optical center. In the latter case it can be recognized immediately that motion stereo will not work for this particular motion. Hence, if 3D analysis is an issue, an observer centered specification of rotation is preferable. For other issues, e.g. the decomposition of motion into "natural" components, other reference systems may be useful.

ILLUMINATION

Time varying illumination gives rise to time varying imagery - a fact which has not enjoyed much attention so far (CORNELIUS and KANADE 83 is an exception). Yet most real-life scenes show changes due to illumination, e.g. darkening from clouds or shadows cast by moving objects. It is not clear what optical flow should be assigned to such phenomena. For most purposes one would probably want to separate illumination effects from physical motion. None of the procedures for optical flow computation does this, however. There appears to be a tacit agreement on leaving illumination problems to higher-level processing. On the other hand, most of the higher-level algorithms proposed for this matter are severely impaired if fed with such optical flow data. This area surely deserves further research.

3. COMPUTING OPTICAL FLOW

In this section I shall try to point out the major issues in computing optical flow as raised by the approaches taken so far.

QUANTIZATION

If one considers quantization part of the computational procedure, the first point in order concerns the temporal and spatial resolution required for a particular approach. Very often, interimage displacements must not exceed a certain amount if the algorithm is to perform well. For example, in gradient-based optical flow algorithms (e.g. HORN and SCHUNCK 81) the displacements must be small enough to permit a linear approximation of the image intensity function over that distance. Token-based schemes, i.e. algorithms which track prominent image elements may be devised to cope with much larger displacements, depending on the correspondence computation. Some algorithms may deteriorate earlier than others if displacements become too large, e.g. Ullman's minimal mapping (ULLMAN 79) as opposed to Radig 's structural matching (RADIG et al. 80).

It is useful to distinguish two phases of optical flow computation: local displacement computation and motion integration. In the first phase, local evidence is evaluated to obtain local displacement information. In the second phase, global measures or constraints are applied to yield a globally consistent flow field. The remainder of this section deals with the two phases in order.

LOCAL DISPLACEMENT COMPUTATION

There are essentially two ways of computing displacements from local information. The first way is directly based on the intensity array. Displacement information is obtained by explaining a temporal intensity change in terms of the displacement of a sloping intensity function. The original version of this approach (due to HORN and SCHUNCK 81) determines only one component of each displacement vector locally. The other component follows from applying a global smoothness constraint. The local ambiguity is not inherent to this approach, however, as can be seen from NAGEL and ENKELMANN 82 and NAGEL 83.

In the second type of approach prominent image features are extracted as a basis for optical flow computation. Further processing depends on whether these features are edge-like or point-like. Edge displacements can usually only be determined with one remaining degree of freedom, hence global criteria must be brought to bear for disambiguation (HILDRETH 83). Point features give rise to the correspondence problem which is disambiguation of displacements in another disguise. It is interesting to observe that in spite of the superficial variety of point and edge operators, the idea is generally the same: edges are locus curves of steepest intensity slope, points are high curvature locations on edges. Nagel's contribution to this workshop is a valuable step towards a unified treatment of edges and points.

MOTION INTEGRATION

What criteria are used to obtain a globally consistent flow field from local measurements? The criterion most frequently encountered is 2D-spatial and temporal smoothness which can be achieved, for example, by minimizing the sum of the squared local velocity differences (HORN and SCHUNCK 81). It is clear that any criterion based on properties of the uninterpreted image will yield a flow field which is in general physically inconsistent, i.e. flow vectors do not always conform with the displacements of physical surface elements. This is particularly evident at object boundaries where the smoothness criterion is not even approximately true. One can cope with this to some extent by keeping track of the smoothing error. In areas where it is large, boundaries may be suspected and a different approach can be taken (CORNELIUS and KANADE 83). Note also the weighting matrix proposed in NAGEL 83. It enforces the smoothness criterion in areas of high ambiguity while preserving the local measurements if based on prominent features (corner points). A similar idea has been followed in YACHIDA 81 where flow vectors are interpolated between prominent point displacements.

It is interesting to compare the computational approaches with psychophysical evidence, e.g. as presented in ADELSON and MOVSHON 83. Within certain limits humans tend to resolve local motion ambiguity following the rule of "common fate". This is in agreement with the smoothness criterion discussed above.

In cases of restricted motion it may be possible to bring to bear a much stronger consistency criterion. If, for example, observer and object translate w.r.t. each other (without rotation), the optical flow field must exhibit a focus of expansion (FOE). Hence all displacement vectors associated with that object must intersect in one point. It should be kept in mind, however, that this is only true for the displacement of physical surface elements. Errors from tracking equal intensity values instead of surface elements may make it difficult to find a FOE.

4. EXPLOITING OPTICAL FLOW

In Fig. 1 two directions for exploiting optical flow are indicated: segmentation and 3D analysis. Both are major areas of research which cannot be adequately covered in this introductory survey. Suffice it to point out some important aspects of these tasks.

The key element of motion-based segmentation is flow discontinuity at potential object boundaries or equivalently homogeneity within boundaries. This is about as valid as the analogous assumption in intensity-based segmentation. Both types of errors - missed and false boundaries - are bound to occur in situations of unrestricted motion. Yet for a subsequent 3D analysis most computational procedures require single-object data. NEUMANN 80 investigates a way of grouping flow vectors into independently moving objects based on 3D interpretability. The Hough-transform approach of O'ROURKE 81 is based on the same idea but may be more suitable for a dense flow field.

Over the last few years the mathematics of motion stereo have been clarified considerably. This is also exemplified by a contribution to this workshop (YEN and HUANG 83) which offers a geometrical interpretation for a formal approach developed earlier. Most theoretical results are based on a certain number of points observed in a certain number of images, e.g. 5 points in 2 images. As such the analysis is not well adapted to a dense optical flow field. Overconstrained sets of equations have been formulated to incorporate a larger number of points. The resulting optimization problem is nontrivial due to a potentially high noise sensitivity (FANG and HUANG 83). When extending 3D analysis to deal with a dense flow field instead of isolated points, one has to bear in mind how the field has been computed. One cannot gain much by incorporating flow vectors which have resulted from interpolation.

Theoretical formulations differ depending on whether a central or parallel projection model is adopted. Proponents of parallel projection usually enjoy less complex mathematics and argue that motion stereo should not be based on perspective effects anyway since these effects may be very small for distant objects. In view of the small number of experiments which have been carried out with real-life data and the scarcety of theoretical results pertaining to noise sensitivity, it is difficult to see decisive advantages with one or the other model.

This concludes the brief survey of approaches to computing and exploiting optical flow. Only a small fraction of the relevant work has been cited. The idea was to give at least one reference - preferably a paper presented at the workshop - for each issue addressed in the discussion. It is hoped that the problems which have been treated so far as well as some open questions have become apparent.

5. REFERENCES

IWMPR = International Workshop on Motion: Representation and Perception, Toronto, Canada, April 3-5, 1983

Adelson and Movshon 83
The Perception of Coherent Motion in Two-Dimensional Patterns
E.H. Adelson, J.A. Movshon
IWMRP 11-16 (1983)

Bonde 79
Untersuchungen zur dreidimensionalen Modellierung bewegter Objekte durch Analyse von Formveraenderungen der Objektbilder in TV-Aufnahmefolgen
T. Bonde
Diplomarbeit, FB Informatik, Universitaet Hamburg, 1979

Cornelius and Kanade 83
Adapting Optical-Flow to Measure Object Motion in Reflectance and X-Ray Image Sequences
N. Cornelius, T. Kanade
IWMRP 50-58 (1983)

Fang and Huang 83
Some Experiments on Estimating 3-D Motion Parameters of a

Rigid Body from Two Consecutive Image Frames
J.Q. Fang, T.S. Huang
(to appear in IEEE-PAMI)

Gibson 50
The Perception of the Visual World
J.J. Gibson
Houghton Mifflin Co., Boston/MA 1950 reprint by Greenwood Press, Westport/CT, 1974

Hildreth 83
Computing the Velocity Field Along Contours
E. Hildreth
IWMRP 26-32 (1983)

Horn and Schunck 81
Determining Optical Flow
B.K.P. Horn and B.G. Schunck
Artificial Intelligence 17 (1981) 185-203

Jain 83
Complex Logarithmic Mapping and the Focus of Expansion
R. Jain
IWMRP 42-49 (1983)

Jerian and Jain 83
Determining Motion Parameters for Scenes with Translation and Rotation
C. Jerian, R. Jain
IWMRP 71-77 (1983)

Nagel 83
On the Estimation of Dense Displacement Vector Fields from Image Sequences
H.-H. Nagel
IWMRP 59-65 (1983)

Nagel and Enkelmann 82
Investigation of Second Order Greyvalue Variations to Estimate Corner Point Displacements
H.-H. Nagel, W. Enkelmann
ICPR-82 768-773 (1982)

Nagel and Neumann 81
On 3D Reconstruction from Two Perspective Views

H.-H. Nagel and B. Neumann
IJCAI-81, 661-663

Neumann 80
Motion Analysis of Image Sequences for Object Grouping and Reconstruction
B. Neumann
ICPR-80, 1262-1265

O'Rourke 81
Dynamically Quantized Spaces Applied to Motion Analysis
J. O'Rourke
JHU-EE 81-1 (January 1981) Electrical Engineering Department The Johns Hopkins University, Baltimore/MD

Prazdny 80
Egomotion and Relative Depth Maps from Optical Flows
K. Prazdny
Biol. Cyb. Vol. 36, 87-102 (1980)

Radig et al. 80
Matching Symbolic Descriptions for 3-D Reconstruction of Simple Moving Objects
B. Radig, R. Kraasch, W. Zach
ICPR-80 1081-1084 (1980)

Rieger and Lawton 83
Determining the Instantaneous Axis of Translation from Optic Flow Generated by Arbitrary Sensor Motion
J.H. Rieger, D.T. Lawton
IWMRP 33-41 (1983)

Tsai and Huang 81
Estimating Three-Dimensional Motion Parameters of a Rigid Planar Patch, I
R.Y. Tsai, T.S. Huang
IEEE-ASSP-29, 1147-1152 (1981)

Tsai and Huang 82
Estimating Three-Dimensional Motion Parameters of a Rigid Planar Patch, II
R.Y. Tsai, T.S. Huang
IEEE-ASSP-30, 552-534 (1982)

Ullman 79

The Interpretation of Visual Motion
S. Ullman
MIT Press, 1979

Webb and Aggarwal 80
Observing Jointed Objects
J.A. Webb, J.K. Aggarwal
ICPR-80, 1246-1250 (1980)

Yachida 81
Determining Velocity Map by 3-D Iterative Estimation
M. Yachida
IJCAI-81 716-718 (1981)

Yen and Huang 83
Determining 3-D Motion Parameters of a Rigid Body: A Vector-Geometrical Approach
B.L. Yen, T.S. Huang
IWMRP 78-90 (1983)